DOCUMENT RESUME

ED 362 565 TM 020 620

AUTHOR Snow, Wendy M.
TITLE The Effects of Repetitive Assessment on Academic Achievement and Pupil Self-Esteem.
PUB DATE 93
NOTE 49p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Educational Assessment; Educational Policy; Elementary School Students; Elementary School Teachers; Grade 5; Intermediate Grades; *Parent Attitudes; Scores; *Self Esteem; *Student Attitudes; *Teacher Attitudes; Teaching Methods; Test Results; Test Use
IDENTIFIERS Parent Surveys; *Repetition Effects; *Retesting; Student Surveys; Teacher Surveys

ABSTRACT

A strategy that has surfaced as a means of promoting success in the classroom is that of automatic retesting for all children whose scores are unsatisfactory on any administered test. The purpose of this investigation was to determine whether the automatic retesting of every student not scoring above a certain point on unit tests is actually beneficial for the child. Two fifth-grade classes of 29 students in Virginia were chosen for the study. Unit tests and retests administered over 7 months were recorded for science, mathematics, social studies, and English; and students were asked to respond to a brief survey asking their feelings about the retest policy. Sixteen teachers and parents of 20 children also responded to surveys about the policy. Average test scores were increased when retesting with the original occurred, and scores in most areas were increased by similar, but not identical, retests. Survey results suggest that parents, teachers, and students benefit from retesting, although some parents have concerns about the fairness of teacher retesting policies. Implications for classroom practice and further research are discussed. Three figures and one table present study findings. Three appendixes contain the surveys. (Contains 6 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM

ED 362 565

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Wendy M. Snow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Effects of Repetitive Assessment on Academic Achievement and Pupil Self-Esteem

Wendy M. Snow
Curry School of Education
University of Virginia

Running Head: EFFECTS OF REPETITIVE ASSESSMENT

TM 020620

The Effects of Repetitive Assessment on Academic Achievement and Pupil Self-Esteem

CHAPTER I

Throughout history the field of education has been continually developing and growing until it has reached its current state. As one looks back in retrospect to examine this evolution, it becomes apparent that this discipline has been filled with various cycles of trends which were all thought at one point to improve learning for our children. For example, at one time there was a greater shift towards more rigorous standards for all children regardless of their academic levels. While educators assumed a more stern and regimental approach to teaching, students were expected to commit themselves to the completion of a predetermined course of study within a set period of time. Nonetheless, even as powerful and lasting as this movement was, it too appears to have floundered somewhat throughout time. Therefore, the search for an ideal means of educating our children is once again

unfolding.

In opposition to the strict stringent approach, much literature is now beginning to support the fact that children need time to progress through school at a pace which is appropriate not only to their chronological age, but to their developmental level as well. Thus, it has been suggested that children should not be forced to complete certain skills within pre-set time limits, such as specific grade-level criterion, but rather should be educated according to their readiness for that particular subject matter. Perhaps one of the greatest manifestations of this movement is the belief that children should be given every opportunity to succeed in the classroom. Consequently, teachers are expected to be more understanding of the differing needs of their students. It is not enough to find and utilize all means necessary to equip each child for academic achievement: Teachers today are additionally faced with the task of empowering students for personal success through higher levels of self-esteem. The agreement among educators in favor of this promotion of individual fulfillment is resounding;

nevertheless, the manner in which to accomplish this ideal has become an issue of debate among many.

One strategy which has surfaced as a means of promoting success in the classroom is that of automatic retesting for all children who score unsatisfactory on any administered test. The general philosophy behind the theory is that every child who does poorly on a test administered by his/her teacher should be given the opportunity to "try again." In giving second chances to our children, schools are viewed more positively and fairly; therefore, students not only do better in school but they gain a love for learning as well. While in principle this theory sounds relatively accommodating to both students and educators, the particular effects of this ideal remain unknown because of an unusual lack of research in this area of testing.

Thus, the purpose of this investigation is to determine whether automatic retesting of every student who does not score above a certain point on unit tests is actually beneficial for the child. It is my plan to further evaluate whether retesting in the classroom: (1) is an effective usage of educational time,

(2) helps students to improve their overall test scores, (3) is beneficial and to what parties (students, teachers, parents), (4) is a positive strategy to use for all students, and (5) is perceived to be beneficial by students, teachers, and parents.

The investigation which follows attempts to evaluate some of the above mentioned questions. From a small school in a rural county in central Virginia, two fifth grade classes of approximately 36 students were chosen for this study. Unit tests administered over a seven month period, including originals and retests, were recorded for the following subjects: science, mathematics, social studies, and English. In an attempt to correlate test scores with student attitudes towards the retesting issue, these students were then asked to respond to a brief survey which questioned their feelings about the policy. Parents and teachers were also asked to complete a survey in relation to their concerns/opinions regarding this issue.

CHAPTER II

Throughout the years it has been an object of debate as to which manner of testing is most beneficial for students and teachers. At one point in American education there was little choice as to which method of testing to use due to a lack of actual writing materials. As a result, almost all tests were administered orally, resulting in a very time-consuming and lengthy means of evaluation. However, as time and technology have progressed, teachers have moved into utilizing the essay and objective test formats which fostered a less complicated method of administering examinations and grading student progress. However, these advancements also brought with them many questions which initially remained unaddressed. For example, how often should students be tested on a particular subject matter and furthermore, how would this new and frequent use of testing going to effect personal achievement in both school and society alike.

Endeavors to answer these questions have remained

relatively untouched until recent decades. However, with the current push for measurable standards of tracking student success in the classroom, a definite shift in the focus of testing is noticeable. On the one hand, testing advocates argue that giving students frequent tests on material enables them to receive more regular feedback from their teachers. Similarly, additional testing serves as a means of increasing not only instructional effectiveness, but it encourages students to study on a more regular basis. On the other hand, researchers have stated that utilizing frequent testing in the classroom takes time away from teachers' much valued and necessary instructional time. For example, if a teacher is forced to give up his/her instructional periods, it is possible that the creativity and ingenuity of that teacher then becomes stifled as he/she is forced to spend more time teaching "to the test." Likewise, students may find it beneficial to focus much of their efforts towards performing well on these endless tests rather than challenging the existing beliefs. This concentration on testing might therefore inhibit enthusiasm for

learning, in both students and teachers, because there is little remaining time to actually experience learning through more creative, hands-on activities. As researchers in the field of testing have begun to address these concerns, a controversy has risen as to the optimum amount of retesting. (Bangert-Drowns, 1991).

One form of retesting which became somewhat popular during the 1980's was that of Mastery Testing. Mastery testing can be described as follows:

> Mastery tests are . . . used at frequent intervals during instruction to evaluate and guide student progress. Students who do not show mastery have their weaknesses diagnosed, receive corrective instruction, and are given new opportunities to show mastery. In most cases, students are not allowed to advance to new material until they show mastery of the objectives covered in an earlier quiz. (Bangert-Drowns, 1991, p. 90)

In comparison to conventional teaching methods, mastery learning provides each individual with the opportunity for success. Since the program is assessed daily on an individual basis, there is no personal competition in the classroom and thus, the class itself is more conducive to learning. Additionally, it is thought that children entering into a conventional classroom at

the bottom of the population distribution have no chance of upward mobility due to class competition. Given the nature of mastery testing, each student is given the amount and kind of individual instruction necessary to attain success in that classroom. Thus, the final result of this form of instruction should be a high level of achievement by all. (Kulik & Kulik, 1987)

While it is important to note that much research shows that programs incorporating mastery testing as one of their features often have positive effects on students, the results do not indicate that frequent testing alone is sufficient for academic achievement (Bangert-Drowns, 1991). On the contrary, mastery testing is a combination of frequent testing along with several other features such as specified feedback on test items and test performance. Thus, it is important to recognize that the literature on mastery testing does not "establish that simply increasing the number of tests in a class will increase student learning." (Bangert-Drowns, 1991)

Although mastery testing is perhaps the most

popular approach utilized in recent times, there are several other testing programs which are currently being employed by educators. Herbert Friedman (1987) offers a method of retesting students in his article "Repeat Examinations in Introductory Statistics Courses." According to Friedman (1987), retesting addresses two major disadvantages in the field of education: "(a) most students find (certain material) anxiety provoking, and (b) the less capable students tend to fall increasingly behind as the course progresses." Thus, it is implied that through the use of retesting, students not only experience less apprehension about their classes, but all students regardless of academic ability are given the opportunity for success.

As stated in Friedman's policy (1987), students were given open-book exams on all original test items. No grades were curved so that those students scoring well on original tests "were not penalized by the others improving their scores on the repeat" and the competitive pressure could be removed. Secondly, any student, regardless of initial score, was permitted to

take the retest with the average of both test scores being recorded. Finally, it was emphasized that the primary purpose of an initial exam was to provide feedback (not grades) for preparation for the repeat test.

As the results indicate, 92% of the repeat exams showed higher grades than on the originals which, "after being averaged with the initial grade, resulted in an increase of just over one letter grade for the exam." (Friedman, 1987) Through surveys, students indicated that they felt less anxiety regarding test-taking, were kept from falling behind, and experienced greater learning because of this retesting policy. It was therefore concluded that:

> Either taking the repeat exam directly led to higher final exam grades, or the more motivated students used the repeat exam option. However, in either case the repeat exams provided support for those students who devoted extra effort to mastering the course material. (Friedman, 1987, p. 22)

While it is important to remark on the benefits found in Friedman's research (1987), there are two issues which remain unaddressed. To begin with, there is no attempt made to identify those students who

simply were more intrinsically motivated to learn regardless of the testing policies administered. As stated above, it is not conclusive whether the repeat exam itself was enough to raise student scores or if the "more motivated" students simply devoted the necessary time for mastering the subject. This issue is significant given that teachers in the elementary school setting are striving continually to foster methods of improving the self-esteem and academic achievement of all students, and particularly the low ability students who may be at risk for later instability. Furthermore, Friedman (1987) admits that "the general similarity of the repeat exams . . . would have made (them) somewhat easier" and may therefore explain the increase in scores.

In an effort to combat this success due to identical repeat exams, Ward Mitchell Cates (1981) established a retesting program in which exams were equivalent (but not identical) to the original tests. In addition, there were fewer retests offered so that a student could retest only once during a given time period and could therefore raise his/her score on only

one original test. As a result, "reducing the number of available retests and increasing the importance of performing well on first" tests appeared often to produce an increase in student effort. All retests were similar in difficulty to the original tests, and they covered the same material; however, test items were not necessarily identical in substance or arrangement. All students were given the option of when to retest, and the highest score of the two tests was always taken.

According to Cates (1981), there are a number of advantages to testing in this manner. To begin, complaints about testing and grading were considerably reduced because students were given the opportunity to retake any tests on which they did not perform satisfactory. It is thought that students under this program take more responsibility for their performance due to the choice involved of when to retake, and the "fact that almost 69% of students do improve on retests seems to quiet those who do not improve." (Cates, 1981) Furthermore, with 88% of students rating the use of retests as "very good" or "excellent," those students

involved with the program appear to be pleased with its outcome.

Again, there are two points worthy of additional note from Cates investigation (1987). To begin with, Cates (1987) suggests that educators should not give the same number of retests as original tests because there is evidence that this practice actually reduces student achievement. The value of this statement is significant for the present investigation because the current policy of the school system under discussion states that retests are to be administered to every student scoring at least below a C on the original. Moreover, Cates (1987) states that giving the same test "will not measure increased mastery of the subject matter but will rather measure memorization" of the material. Since actual learning, not rote memorization, is the mission of our school systems, it only makes sense that teachers implementing a retest policy should consider their objectives of learning for students before giving students the same test twice. Students having received an identical test twice and still not attaining a satisfactory score may feel an

even greater sense of failure, which in essence defeats the entire theory behind any retesting policy.

Because of the documentation that frequent testing alone does not signify achievement gains, it is now important to look at what factors do contribute to successful testing policies. For the purpose of this investigation, it is important to note that the classes observed were not utilizing the mastery testing approach. Instead, two fifth grade classes were observed in which the retest policies varied depending on the teacher, subject matter, time constraints, etc. One teacher often gave completely new tests which contained different questions while the other often corrected the original test and asked students to "fix" any questions that were marked wrong. While keeping in mind the diversity of the testing procedures observed in this study, it is even more important to note that there is very little literature relating to any form of retesting other than that of mastery learning. Furthermore, the sparse information which is available regarding retesting involves mainly students at the collegiate level.

CHAPTER III

Method

Subjects

Students:

Thirty-six fifth grade students in a small rural school in central Virginia participated in this investigation. These students, 16 males and 20 females between the ages of ten and eleven, were divided into two heterogenous groups referred to in this study as CLASS A and CLASS B. CLASS A, which contained 19 students, and CLASS B, 17 students, were grouped in this manner for the duration of the study. The county in which this school is located initiated a testing policy for the current academic year in which all students scoring unsatisfactory on tests are to be given retests on that same material.

Students were administered surveys which asked them to answer questions regarding the retest policy. Since participation in the study was voluntary, seven of the thirty-six students opted not to complete the survey.

Parents:

Twenty parents from the school were selected to complete a survey regarding their feelings toward this retest policy. These parents and may have had more than one child attending this 4-6 grade school. Furthermore, parents ranged in intellect from those possessing an elementary education solely to others holding college degrees. Due to the voluntary participation of the surveys, only 12 parents (60%) chose to take part in this effort and to return the questionnaires to the school. Finally, because of the anonymity of the survey, little else was known about the recipients of the parental survey.

Teachers:

Sixteen teachers (the entire licensed faculty at this small school), including two special education and two physical education instructors, were asked to complete a survey for this study. Of the fifteen female and one male educators surveyed, all hold valid Virginia teaching licensing with one teacher possessing a Masters degree in Reading Diagnosis. Once again, due

to the voluntary nature of the surveys administered, 7 teachers (47%) completed and returned their surveys for this investigation.

Measures

Because of the differing methods of retesting utilized in CLASS A and CLASS B, it is useful to discuss these two groups in isolation. Furthermore, it is purposeful to separate the measures used in this study according to two categories: (1) quantitative data (test scores) and (2) qualitative data (surveys).

Test Scores

CLASS A:

This experiment was conducted using all unit test scores recorded during a seven month period of time. Unit test scores were recorded for four subjects during this time period: social studies, science, mathematics, and English. During the first four months of observation, students in this condition scoring below a 77 were administered retests automatically. Additionally, any student scoring above a 77 wishing to

improve his/her score was permitted to take the retest as well. Although the retest covered the same material, this evaluative item was not identical to the original test administered. Finally, the instructor of CLASS A recorded the highest of both test scores as the final unit test grade.

After four months of implementing this type of retesting, the instructor of this condition dropped from usage the retest policy entirely. Therefore, test scores recorded during this period of time were of the original test items only. In other words, students in CLASS A were at one time administered retests; however, currently students are given one and only one unit test.

CLASS B:

As seen in CLASS A, unit test scores were recorded for a seven month time period across four subjects: social studies, science, mathematics, and English. The instructor automatically administered retests (which were similar but not identical to the original tests) to all students scoring a D or F, or below a 77. As

mentioned above, any student requesting to take a retest was also permitted to do so, and the highest of both test scores was recorded in all instances. However, CLASS B also changed its methods of retesting during this seven month period of time.

Finding this method of retesting somewhat time-consuming, CLASS B altered its retest policy to the following: All original tests were marked for correct answers and students were given a numerical grade. Afterwards, students were simply returned their original tests and instructed to "fix" any problems marked as wrong. Thus, after having adjusted all incorrect answers, the teacher then re-graded the original corrected test and recorded the new, higher score.

Surveys

Students:

In an effort to obtain student opinion regarding the retesting policy, a one page survey was administered to both CLASS A and CLASS B (**Appendix A**). Of the thirty-six students involved in the study, twenty-nine voluntarily agreed to complete the thirteen

question survey. In order to ascertain whether retesting actually improved self-esteem, students were asked to state their views concerning the policy.

Parents:

The parents of twenty students were issued a cover letter plus a one page survey asking them to complete thirteen questions regarding their feelings towards the retest policy (**Appendix B**). These surveys were sent home from school via the children. Parents were given a one week period of time to complete and return surveys to their respective teachers. The parent survey not only assessed beliefs concerning the policy itself, but it also questioned parental views on the fairness of teacher implementation of such retesting.

Teachers:

Sixteen teachers received surveys in their school mailboxes regarding the retest policy in their county. This survey, consisting of a cover letter and two additional pages of questions, assessed teacher beliefs about retesting in relation to educational time,

student success, teacher demands, etc. (**Appendix C**). Teachers were given a one week period of time to complete and return surveys to the school office.

Design

In the present study, a between subjects design was selected. The independent variable was the type of retest policy administered, and there were three levels in this variable: (1) No retest, (2) Similar but not identical retest, and (3) Retest using original test. Those subjects participating in the "No retest" condition were given unit tests in social studies, science, mathematics, and English and were not offered a repeat examination. The unit test scores for participants were then recorded. Subjects in the "Similar but not identical retest" condition were administered an initial test on material, and all students scoring below a 77 were required to take a similar (but not identical) retests on the same information. Furthermore, any student scoring above 77 wishing to take the retest was permitted to do so. Once again, all unit test scores were recorded for the

classes of social studies, science, mathematics, and English with the highest of the test/retest being utilized. This procedure was followed for the "Retest with original" as well. However, in this condition students were given an original test which was corrected by the teacher. Afterwards, the teacher returned the test and had students simply "fix" those answers marked wrong. After corrections were made, the teacher then re-scored the test and recorded the highest grade. The dependent measure in this investigation was the average scores of all students per discipline (social studies, science, mathematics, and English) with regard to retesting procedure. Thus, after all data was collected, a mean test score for social studies, science, mathematics, and English was calculated for each of the three testing condition.

Surveys were utilized as a means of receiving opinions regarding the retest policy from teachers, parents, and students. In all surveys, individuals were asked to score their answers by circling a number one through five with 1=STRONGLY DISAGREE and 5=STRONGLY AGREE. Surveys were separated into piles

according to classification of either teacher, student, or parent. Afterwards, questions per survey were looked at individually so that it could be determined how many subjects answered 1, answered 2, answered 3, etc. Finally, all outstanding trends were recorded and compared among the three survey groups (**Table I**).

Analysis

Test scores were collected for each of the students in CLASS A and CLASS B across the four observed subject areas: science, social studies, mathematics, and English. Once all test scores were collected, test scores for both CLASS A and CLASS B were divided into type of retesting method utilized. For CLASS A, mean test scores were calculated for both types of testing policies across the four subject areas. As shown in Figure 1.1, these mean test scores were then compared per discipline by the type of retesting policy used.

Insert Figure 1.1 about here

In CLASS B, test scores were divided into subject areas as well as type of retesting policy used. Then mean scores were calculated first for the "retest with similar" condition for each of the four disciplines. This procedure was again followed for the "retest with original" condition. After having collected the mean test scores for both conditions, these testing policies were compared according to the subject area (**Figure 1.2**).

Insert Figure 1.2 about here

After having compared the testing policies in each class, the differing types of retesting were then compared in Figure 1.3. The mean test scores from the "no retest" condition in CLASS A was compared to the mean test scores from the "retest with original" condition in CLASS B for the four subject areas.

Insert Figure 1.3 about here

Turning to the results of the surveys, the

experimenter separated all surveys into one of the three types of respondents: students, teachers, and parents. Beginning with the student surveys, each of the 13 questions were observed separately so that it could be determined how many subjects answered 1, answered 2, and so on. Once these results had been tabulated, a percentage of students somewhat agreeing to strongly agreeing with the question was determined. This method of analyzing the surveys was continued for the remaining two parties, parents and teachers. Then once all percentages had been tabulated, the experimenter examined the results to ascertain any predominant trends across the three groups. These trends were then compared and displayed in **Table 1**.

A discussion will ensue in the following chapter regarding the various retesting policies mentioned in this study. The question of whether retesting actually helps students to improve overall test scores will be addressed by comparing mean test scores across testing methods and subject disciplines. Furthermore, the perceptions of these policies will be observed with regard to student, parent, and teacher beliefs.

CHAPTER IV

Results

In the present investigation, it is inferred from the results that retesting is beneficial in raising overall test scores across social studies, science, mathematics, and English. When focusing on CLASS A (which utilized "Similar but not identical" retests for four months and then went to a "No retest" policy for the remaining three month period), it is clear that retesting raised mean scores for each subject area considered except for social studies (**Figure 1.1**). CLASS B found similar results such that students in the "Retest with original" condition scored higher overall than when administering a "Similar but not identical" retest. This effect was found for all area and disciplines observed except for science (**Figure 1.2**). Furthermore, when comparing average test scores across the four disciplines in relation to retesting policy, it is interesting to note that "Retesting with the Original" yielded much higher scores in social studies and mathematics than did "Retesting with a different

measure" (**Figure 1.3**). On the contrary, the mean score for science was higher when "Retesting with a different measure" while the English means were exactly the same regardless of the type of retesting policy utilized.

Retesting Vs. No Retesting:

In order to determine whether retesting increases overall test scores, it is beneficial to focus this portion of the results on CLASS A which consisted of two conditions: (1) Similar but not identical retests and (2) No retests issued. In most cases it does appear that retesting raises overall test scores except for the subject of social studies (**Figure 1.1**). Three subjects appeared to have an increase in overall averages when administered a similar, but not identical retest. For example, through the use of retesting average scores were raised as follows: Science: 2.0 points, English: 5.1 points, and Mathematics: 1.0 points. However, students scored an average of 1.3 points higher on social studies unit tests when not offered a retest.

Type of Retest:

To determine which method of retesting is most beneficial for students, it is important to focus this portion of the results on CLASS B. Using the standard form of retesting for several months (retesting with a similar, but not identical test), the instructor of CLASS B opted to change testing methods such that students were simply retested by having them correct or "fix" any wrong items on their original test papers. Using the method of correcting the original tests, students scored higher on the average per subject in every class except for science (**Figure 1.2**). Students scored higher on average when given the original test to correct in the following subjects: Social studies: 7.0 points, English: 3.0 points, and Mathematics: 7.0 points. On the contrary, students in science class scored an average of 0.5 points higher when given a similar, but not identical retest measure.

Effective Use of Educational Time:

It is apparent that in most cases the overall scores are raised when a retest policy is in place.

However, there are great concerns about the effectiveness of using this policy. For example, in reviewing the results of CLASS A, it is evident that although test scores are generally raised, this increased average is not significantly different than that average calculated from the no retesting condition. Furthermore, while looking at CLASS B, even though there are significant gains in test scores after giving back the original corrected test, much literature supports the fact that this increase in test scores is due to mere rote memorization of information rather than true learning. (Bangert-Drowns, 1991) As a result, we are faced with the question of whether we are testing simply for the purpose of showing higher test scores or are we using testing as a means of enhancing actual learning. Thus, to retest after every single evaluative item is not an effective use of educational time due to the fact that (1) the increase in test scores is not significant, and (2) certain forms of retesting may actually stifle true learning by forcing students to simply memorize information.

Parties Benefitted:

When asked on a survey if retesting was beneficial for children, 75% of those parents surveyed agreed that the retesting policy was good. This number is high in comparison to only 66% of student agreement and 57% of teacher agreement. Furthermore, while 92% of all parents felt that retesting improved their child's overall grades, only 66% of students felt similarly.

Positive for All Students:

When asked whether retesting was a positive experience for children, 100% of teachers surveys agreed that it was. In comparison, only slightly more than one out of two students felt retesting was actually positive. It is also important to note that while 100% of teachers surveyed felt that making a child take a retest did not humiliate the student, approximately 1 out of every 4 students was embarrassed in being singled-out in front of their peers and having to take a retest **(Table 1)**.

CHAPTER V

Discussion

I. Analysis of Results

A. Test Scores

In determining whether retesting is actually beneficial for students, it must be stated that average test scores across the four observed subjects were increased when retesting policies were in place. In CLASS A where there was a similar but not identical retest offered, student test scores were raised somewhat in science and mathematics, and particularly elevated for English (an increase of 5.5 points). However, students did not appear to benefit from this method of retesting in social studies where average test scores actually fell when given retests (1.5 points). Therefore, it can be stated that in most cases average test scores were raised by administering similar but not identical retests as opposed to offering no retest.

When determining which type of retesting policy to implement in the classroom, it becomes useful to

compare average test scores from the two retesting policies: (1) Similar but not identical retest, and (2) Retest with original. It is evident from the averages that retesting with the original test raises test scores particularly in social studies and mathematics, 9.0 points and 7.0 points respectively (**Figure 1.3**). However, in science it appears that students often score better when given a similar but not identical retest (4.0 points). Finally, while students in English appear to benefit greatly from having a retest available to them, it seems they score the same overall regardless of retesting method administered.

B. Surveys

The results from the survey suggest that parents, teachers, and students benefit from retesting. While 92% of parents feel that it improves students' grades, only 50% of parents feel that their child's teacher is consistent and fair in administering retests. This is quite a contrast to teachers who unanimously (100%) feel that they are both consistent and fair in implementing retest policies. Furthermore, while 100%

of teachers feel that they make taking retest a positive experience for children, only 52% of students actually see it as positive with approximately one in four students viewing it as an embarrassment (**Table 1**). Thus, it is clear that there is some discrepancy between parents, teachers, and students as to the actual effects of the retest policy.

II. Implications for Retesting Issue

The results of this study indicate that student scores are generally raised when offered retests. Furthermore, many of the opinions of parents, teachers, and students regarding this policy appear to be positive. However, it must be stressed that approximately 25% of all students undergoing retesting feel that others see them as perhaps a failure and thus, they are embarrassed by having to take retests. Perhaps teachers should give further consideration to the discrepancy in beliefs on whether the child should have the option of when to retest in <u>all</u> instances. That is, while only 29% of teachers feel that it should be the child's ultimate decision of when to retest, 70% of students feel that it should be their option in all

circumstances. Furthermore, although average test scores were raised, many of the elevations were not significant. Thus, one is lead to question whether retesting is a wise use of educational time. While there are those who feel that raising test scores, no matter how minimal, makes retesting worthy, others argue that the time necessary for preparing retests could be more wisely spent in the classroom.

III. Limitations of Study

Due to the small sample of subjects, it is possible that generalities regarding retesting are limited. At the onset of the study, CLASS A and CLASS B were implementing very similar retest policies; however, as time progressed both teachers became dissatisfied and changed testing methods. This change in policy further added to design problems such that the two classes had to be studied in isolation. Additionally, in an effort to determine which types of students are embarrassed by taking retests, it may have been useful to mark the surveys so that they could be compared with individual students test scores. This procedure would also have been helpful in identifying

which types of students actually benefit from being offered a retest.

IV. Implications for Further Research

Given that many school systems are currently looking to methods retesting for improving student scores and self-esteem, it is necessary that more literature focus on this issue. As mentioned earlier, there is a lack of research in this area at present and thus, retesting in the classroom is evolving with very little basis. Because of the time restraints of the present study, it was not possible to determine which type of student--whether of high, medium, or low ability--benefitted most from retesting. Furthermore, are there certain content areas which appear to benefit most from having retests available to students. For example, why is it that students appear to do much better in English when administered retests? And finally, it is important to consider which types of retesting are most beneficial for learning. If the school measures actual learning based on test scores, it is perhaps most advantageous to use the "retesting with original" policy regardless of the fact that much

research states that this merely promotes rote memorization. On the contrary, if a school system attempts to improve actual learning through retesting, there needs to be more sufficient data to link the particular methods of retesting to this ideal. Therefore, a great deal of research is needed in this area given that school systems are accepting retest policies in greater numbers.

References

Bangert-Drowns, R. L. (1991). Effects of frequent classroom testing. Journal of Educational Research, 85, 89-99.

Cates, W. M. (1984). Retesting: a logical alternative in college instruction. Improving College and University Teaching, 32, 99-103.

Friedman, H. (1987). Repeat examinations in introductory statistics courses. Teaching of Psychology, 14, 20-23.

Kirst, M. W. (1991). Interview on assessment with James Popham. Educational Researcher, 20, 24-27.

Kulik, C. L. and Kulik, J. A. (1987). Mastery testing and student learning: a meta-analysis. Journal of Educational Technology Systems, 15, 325-345.

Smith, W. J. (1984). Minimizing teacher concerns about retesting in mastery learning. NASSP Bulletin, 68, 108-112.

Figure 1.1

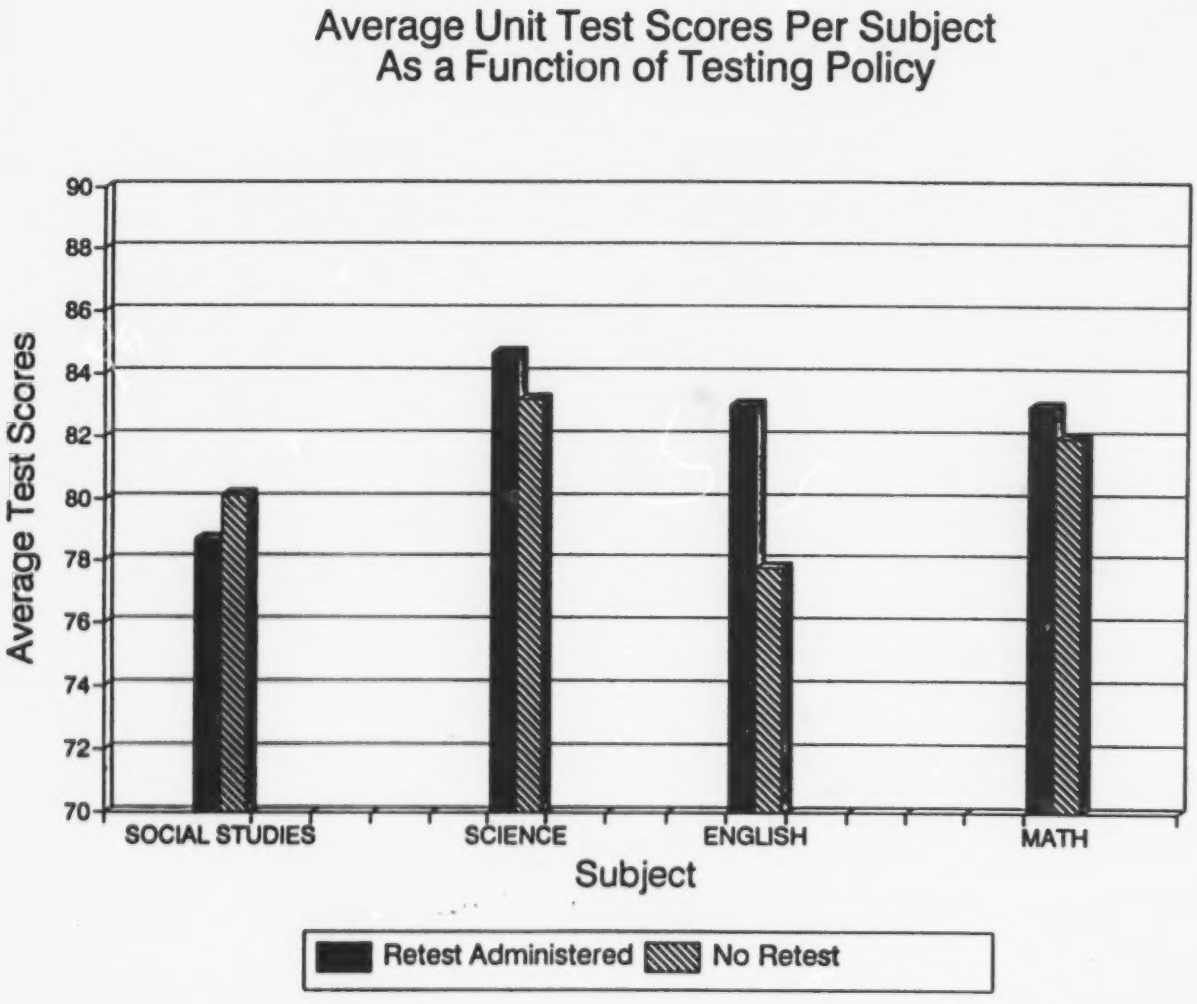
Average Unit Test Scores Per Subject
As a Function of Testing Policy
Average Test Scores
90
88
86
84
82
80
78
76
74
72
70
SOCIAL STUDIES
SCIENCE
ENGLISH
MATH
Subject
Retest Administered
No Retest

Figure 1.2

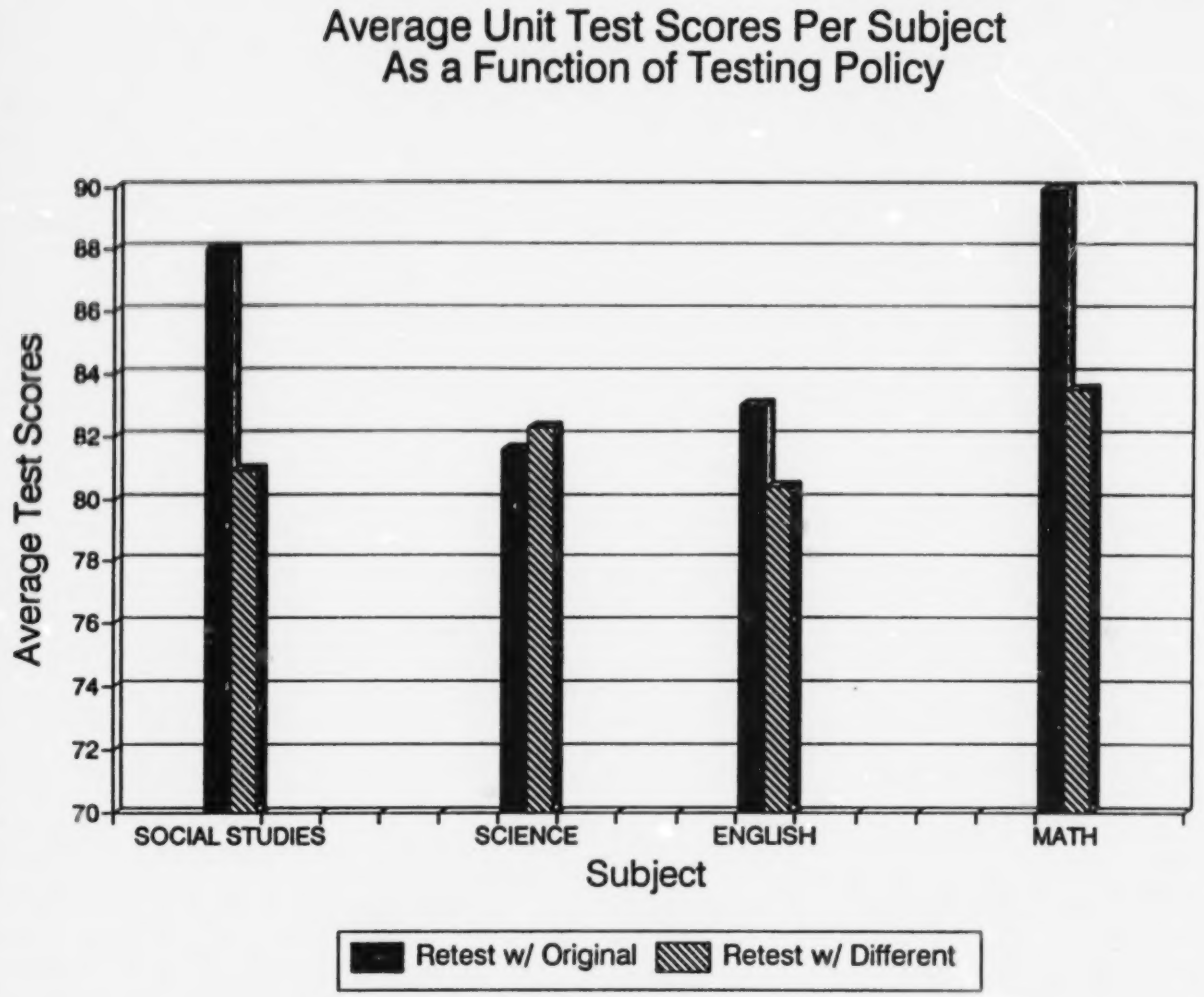
Average Unit Test Scores Per Subject
As a Function of Testing Policy
Average Test Scores
90
88
86
84
82
80
78
76
74
72
70
SOCIAL STUDIES
SCIENCE
ENGLISH
MATH
Subject
Retest w/ Original
Retest w/ Different

Figure 1.3

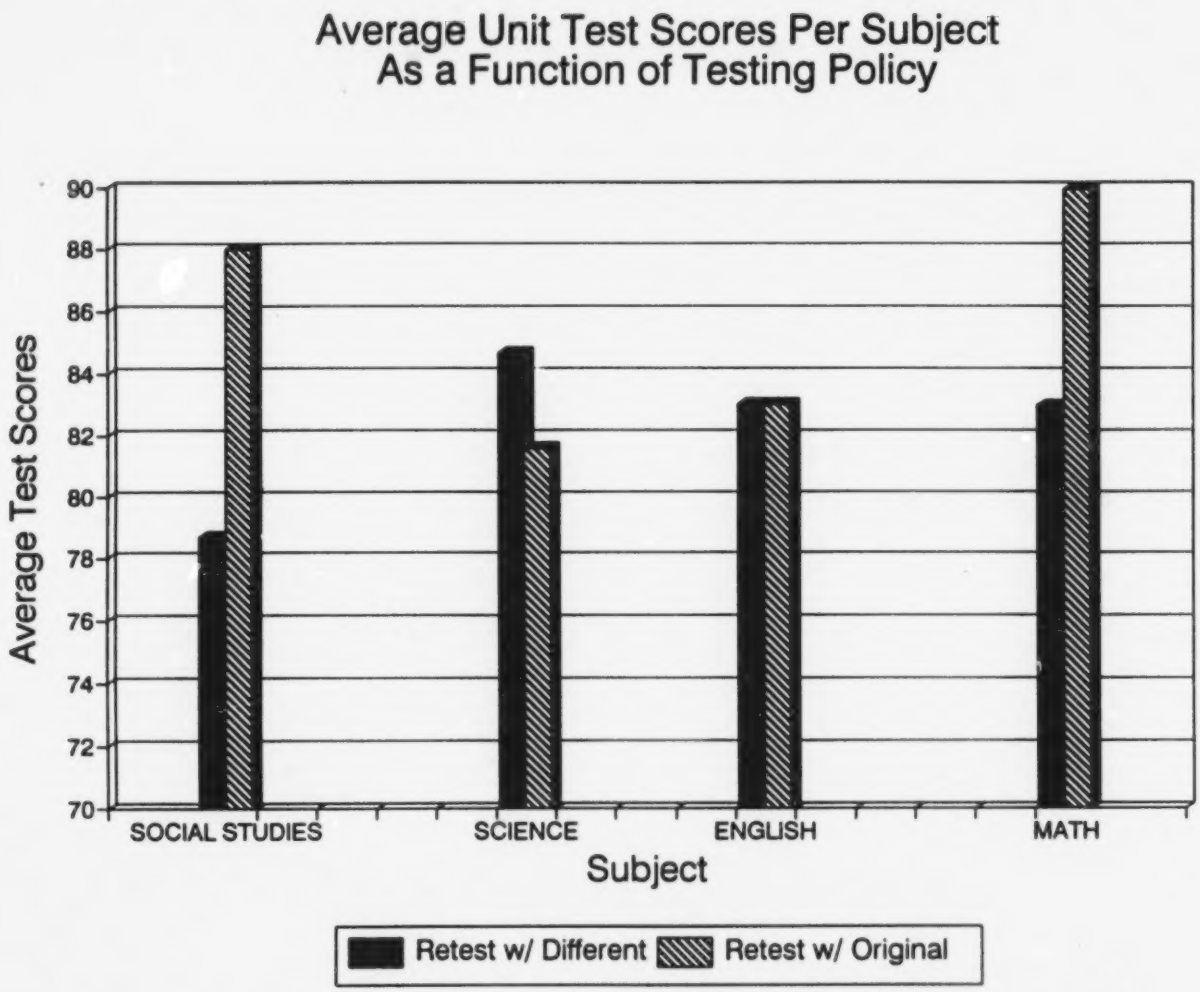
Average Unit Test Scores Per Subject
As a Function of Testing Policy
Average Test Scores
90
88
86
84
82
80
78
76
74
72
70
SOCIAL STUDIES
SCIENCE
ENGLISH
MATH
Subject
Retest w/ Different
Retest w/ Original

Table 1

Analysis of Surveys

Percentage of Subjects
Somewhat Agreeing to Strongly Agreeing
with Regard to Retesting

REMARKS	TEACHERS	STUDENTS	PARENTS
IMPROVES OVERALL GRADES	71%	66%	92%
TEACHER MAKES IT POSITIVE	100%	52%	83%
SEEN AS A PRIVILEGE, NOT REWARD	83%	62%	83%
DOES NOT EMBARRASS STUDENTS	100%	76%	75%
ONLY RETEST WHEN SCORE BELOW C	86%	59%	75%
BENEFICIAL FOR CHILD OVERALL	57%	66%	75%
TEACHER IS CONSISTENT AND FAIR	100%	66%	50%
RETEST IS HARDER THAN ORIGINAL	0%	41%	50%
ALWAYS CHILD'S OPTION TO RETEST	29%	70%	42%

Appendix A

Please answer the following questions by circling the numbers. Your answers on this survey are strictly confidential and will remain anonymous.

Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
1	2	3	4	5

1. *I sometimes do not study for the original test because I know I will have a retest available.*

   1 2 3 4 5

2. *The retest policy is beneficial for me as a student because it improves my grades.*

   1 2 3 4 5

3. *It should be my option as to whether or not I want to take a retest.*

   1 2 3 4 5

4. *My teacher is consistent in her retest policies.*

   1 2 3 4 5

5. *My being able to take a retest is a privilege.*

   1 2 3 4 5

6. *I generally do better on the first test.*

   1 2 3 4 5

7. *Getting to take retests has improved my scores on the whole.*

   1 2 3 4 5

8. *Teachers do not spend much class time in giving retests.*

   1 2 3 4 5

9. *I am not embarrassed when I have to take a retest.*

   1 2 3 4 5

10. *I don't have to miss class, fall behind, miss new material, etc. when I take retests.*

    1 2 3 4 5

11. *Retests are generally harder than the original tests.*

    1 2 3 4 5

12. *My teacher makes taking retests a positive experience.*

    1 2 3 4 5

13. *I should only have to repeat a test when I have scored below a C on the original.*

    1 2 3 4 5

Appendix B

February 28, 1993

Dear Parents,

As you may already be aware, several of the teachers at Lovingston Elementary are offering their students the option of taking a re-test on material which was not passed on the original test. For example, if a student were to complete a unit on Insects, take the test, and then receive an unsatisfactory grade on that test, he or she may be given the opportunity to take a second test on Insects to improve the original score.

Attached you will find a survey which asks you to list your opinions about this retest policy. In an effort to discover its effectiveness, I have chosen to investigate the benefits of this policy and to record the findings in my thesis paper at the University of Virginia. Your answers will be completely anonymous and confidential and will in no way be presented to your child's classroom teacher, principal, community, etc. Therefore, please take the time to give each question careful consideration, and I thank you for your participation in this survey.

Sincerely

Wendy M. Snow
University of Virginia,
Curry School of Education

Please answer the following questions by circling the numbers. Your answers on this survey are strictly confidential and will remain anonymous.

Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
1	2	3	4	5

1. *The policy of automatic retesting in Nelson County is beneficial for my child.*

   1 2 3 4 5

2. *Children should only repeat a test when they score below a C on the original.*

   1 2 3 4 5

3. *Children should be given the option of whether to repeat a test no matter what they scored on the first test.*

   1 2 3 4 5

4. *Teachers are consistent in how they administer their retest policies across subjects.*

   1 2 3 4 5

5. *My child views the opportunity to retake a test as a privilege and a means of bettering his grade.*

   1 2 3 4 5

6. *My child does well in general on the original tests.*

   1 2 3 4 5

7. *My child's scores generally improve on the retest.*

   1 2 3 4 5

8. *The retest policy has given my child a chance to improve his overall grades in school.*

   1 2 3 4 5

9. *Teachers' educational time is well spent in administering retests to students.*

   1 2 3 4 5

10. *My child is embarrassed by having to take retests because he is singled out from the others students.*

    1 2 3 4 5

11. *My child does not miss a great deal of class time, fall behind in new material, etc. when taking retest.*

    1 2 3 4 5

12. *My child's teacher makes taking retests a positive experience.*

    1 2 3 4 5

13. *My child generally finds the retest harder than the original.*

    1 2 3 4 5

Appendix C

February 28, 1993

Dear Teachers,

Attached you will find a survey which asks you several questions concerning your feelings towards the retest policy. As a student teacher here at Lovingston, I had several questions and concerns regarding how the policy should be implemented, and therefore, have decided to do my thesis project on this issue. It is my plan to evaluate whether administering retests in the classroom are (1) an efficient use of educational time, (2) helping students to improve their scores, (3) viewed positively by students, parents, and teachers, (4) actually beneficial and to what parties, and (5) the conditions under which retesting is positive for students.

Therefore, I would appreciate your filling out the attached survey so that I might get a feel for your views on this subject. Your answers will remain anonymous and confidential (so don't sign your name to the survey!), and will be in no way shown to parents, students, or other staff members. There is an envelope in the office for you to leave the surveys once completed.

Thank you so much for your cooperation in this study.

Sincerely,

Wendy M. Snow

Wendy M. Snow

Please answer the following questions by circling the numbers. Your answers on this survey are strictly confidential and will remain anonymous.

Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
1	2	3	4	5

1. *The retest policy in Nelson County is beneficial for my students because it raises their scores.*

   1 2 3 4 5

2. *The retest policy is useful for me as a teacher.*

   1 2 3 4 5

3. *It should be the student's option in all circumstances whether to retake a test.*

   1 2 3 4 5

4. *My retest policies are consistent and fair.*

   1 2 3 4 5

5. *I make taking a retest seem as a privilege, not a reward or embarrassment.*

   1 2 3 4 5

6. *My students generally do better on the original test.*

   1 2 3 4 5

7. *Using retests has improved classroom scores on the whole.*

   1 2 3 4 5

8. *I do not have to spend much additional class time to administer retests.*

   1 2 3 4 5

9. *My students are not embarrassed by having to take retests.*

   1 2 3 4 5

10. *My students do not miss class time, fall behind, miss new material, etc. when taking retests.*

    1 2 3 4 5

11. *I make taking a retest a positive experience.*

    1 2 3 4 5

12. *I make the retest harder than the original.*

    1 2 3 4 5

13. *Only those students scoring below a C should have to take the retest.*

    1 2 3 4 5

14. *I find that certain students are always having to take the retests.*

1 2 3 4 5

15. *I would keep the retest policy in place in future years.*

1 2 3 4 5

16. *Describe how you grade the original and retests. For example, do you take the highest score or average the two grades?*

17. *What is your standard retest policy, such as who has to take the retest?*

18. *Any additional comments which may be helpful:*